

ISSN 0825-4591

PROCEEDINGS

ENVIRONMENTAL RESEARCH:
1989 TECHNOLOGY TRANSFER CONFERENCE

November 20 & 21, 1989

Royal York Hotel

Toronto, Ontario

VOLUME I

FEATURE PRESENTATIONS

AIR QUALITY RESEARCH

WATER QUALITY RESEARCH

Sponsored by
Research and Technology Branch
Environment Ontario
Ontario, Canada

ABWF

Copyright Provisions and Restrictions on Copying:

This Ontario Ministry of the Environment work is protected by Crown copyright (unless otherwise indicated), which is held by the Queen's Printer for Ontario. It may be reproduced for non-commercial purposes if credit is given and Crown copyright is acknowledged.

It may not be reproduced, in all or in part, for any commercial purpose except under a licence from the Queen's Printer for Ontario.

For information on reproducing Government of Ontario works, please contact ServiceOntario Publications at copyright@ontario.ca

Multiple Option Approach to the Risk Assessment For Mixtures: Experience With Polycyclic Aromatic Hydrocarbons (PAH)

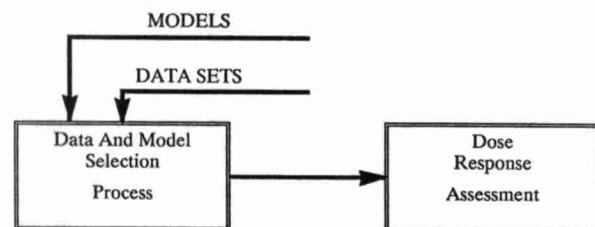
Pavel Muller and Bryan Leece
Ontario Ministry of the Environment
Hazardous Contaminants Co-ordination Branch

1 INTRODUCTION

1.1 Background

Our group is involved in the preparation of a report on the risk to man and other biota from a multimedia exposure to PAH-containing complex mixtures. The approach described below was developed to assist our group in this endeavor. The database for PAH is extensive and many different approaches to risk assessment could be attempted. We have opted for the development of a formal process to guide us through the selection of the best alternative(s). It is useful to visualize this process as a "front end" to the risk assessment process, because the process helps in the selection of the most suitable approaches to risk assessment (models) and the data these models use (See fig.1 and sections 1.2 and 1.3).

Fig. 1: Proposed data and model selection process serves as a front end to the risk assessment process.



1.2 Selection of Models

Over the years, a number of risk assessment models have been applied to PAH and PAH-containing complex mixtures. The sections below provide a description of a few of them.

1.2.1 Benzo[a]pyrene (B[a]P) Equivalents Model

Thorslund and Charnley [1] proposed to convert the potency of individual PAH in a complex mixture into B[a]P equivalents. The concentration of PAH in a complex mixture is then expressed in terms of the equivalents of B[a]P. The potency of such a mixture is assumed to be equal to that of B[a]P of the same concentration. The authors refer to this approach as a "comparative potency approach", a name which is unfortunately the same as that Albert et al., 1983 [2] have proposed for a different model. Our Ministry used a model similar to that of Thorslund and Charnley for dioxins and dibenzofurans, using 2,3,7,8-T₄CDD as the "standard" [3].

To apply the B[a]P Equivalents model to a given complex mixture, concentration data for each PAH within the mixture must be available.

1.2.2 B[a]P Equipotency Model

For risk assessment purposes, it is assumed that all PAH in complex mixtures are equal in potency to benzo[a]pyrene. This approach has been proposed by US EPA [4] and applied to complex mixtures [5].

To apply the B[a]P Equipotency Model to a given complex mixture, total PAH or Polycyclic Organic Matter (POM) concentration data is required.

1.2.3 B[a]P Proportional Potency Model

For risk assessment purposes, it is assumed that the potency of a PAH - containing mixture is directly proportional to its benzo[a]pyrene content. Pike and Henderson [6] suggested that 15 ng/m³ of B[a]P be considered as equivalent in its human lung risk to a single cigarette. WHO used the B[a]P Proportional Potency Model for their Drinking Water Guidelines [7] and Air Quality Guidelines for Europe [8].

Application of the B[a]P Proportional Potency Model requires that the concentration of B[a]P in a complex mixture be known.

1.2.4 Comparative Potency Model

For risk assessment purposes, it is assumed that for all PAH-containing complex mixtures, the ratio of the potency of a mixture in man and the potency of that mixture in an appropriate experimental test is a constant C (see equation 1.)

Equation 1.

$$C = \frac{(\text{humanrisk}(\text{mixt.1}))}{(\text{exper.potency}(\text{mixt.1}))} = \frac{(\text{humanrisk}(\text{mixt.2}))}{(\text{exper.potency}(\text{mixt.2}))}$$

From this equation, the human risk for an unknown complex mixture (mixture 2) can easily be derived by the rearrangement of equation 1 (see equation 2).

My thanks to B. Thorpe and Drs. B. Birmingham and J. Lewtas for their comments and support throughout the development of this approach.

This work was performed as a part of development of multimedia standards for PAH - containing mixtures by the Ontario Ministry of the Environment, Hazardous Contaminants Coordination Branch.

Equation 2.

$$\text{humanrisk}(\text{mixt.2}) = C * \text{exper.potency}(\text{mixt.2})$$

Use of the Comparative Potency Model requires knowledge of the concentration of the complex mixture within a given medium (for example, the amount of benzene-soluble fraction extracted from coke oven emissions).

It is possible to use other models in risk assessment of PAH-containing mixtures and some will be discussed later in this document. Each of these models has its advantages and disadvantages. It is difficult to identify the most promising one without assessing the data sets on which these models rely.

1.3 Selection of Data

Many regulatory agencies use a risk assessment process to estimate human risk of non-threshold toxicants such as carcinogens [9],[10]. Despite the popularity of this process, the estimates derived from it are not always reliable. Most efforts have focused on improving the reliability of the extrapolation from experimental animal studies to man [11],[12],[13],[14],[15],[16],[17],[18]. The reliability of the risk estimate is dependent not only on the sophistication of these extrapolations, but also on the quality and general suitability of the data used for the risk assessment.

In order to estimate risk, a collection of diverse data is required, referred to here as a data set. Some data sets require relatively little extrapolation in order to estimate human risk. We will refer to such data as most "appropriate". For example US EPA [19] recommends the use of human data in preference to animal data (see table 1.), presumably because human data require fewer extrapolations in order to be applicable for establishing risk to general population.

Table 1. Appropriateness of data
(In the order of descending "Appropriateness")

o	Human data in preference to animal data
o	Data from the species responding most like human
o	Other species

(USEPA, 1986 Guidelines for the Carcinogen Risk Assessment)

The most appropriate type of data may not be always available or data which is available may be of poor quality. For example, the data set may be incomplete or derived from studies with design flaws or shortcomings in data analysis. Therefore, it is necessary to consider both the quality and the appropriateness of the data, when selecting data sets for risk assessment. The US EPA addressed the issue of data selection in its guidelines for mixtures [20]. It recommends the selection of the most appropriate data set, for

which the data are of adequate quality. This approach places a relatively heavy emphasis on the appropriateness of the data, and less on the quality of the data. This approach may often be reasonable, considering that the errors resulting from extrapolation from less appropriate data are likely to be larger than the errors which arise from using data of low quality. There are however, alternatives to this approach. The most robust risk estimate will be the one based on data sets where there is an optimal balance between the appropriateness and quality of data. The appropriateness and quality evaluations are performed for each data set and the best candidate(s) are selected.

1.4 Relation between the models and data

While sections 1.2 and 1.3 describe the selection of risk assessment models and their attendant data sets separately, the models and the data sets are, in fact, closely related since selection of one directs the selection of the other. Consequently the models and their respective data sets must be evaluated together. In this report, we describe a formal process for evaluation of the models and data sets, the goal of which is to identify the models and data sets which are likely to provide the most reliable estimate of human risk.

1.5 Rationale for the proposed approach

There were three reasons for developing a formal process for selecting models and data sets for subsequent risk assessment:

- The sheer volume of potentially relevant literature makes it impractical to give equal attention to the entire PAH literature. A formal approach helps identify the most relevant information for risk assessment. The selected areas can then be reviewed in greater depth.
- Throughout the risk assessment process, value judgements are made frequently and unavoidably. Our regulatory process is open to public scrutiny, and it is therefore necessary to make value judgements explicit. Our formal approach has a built-in mechanism to make most of the value judgements explicit and the rationale for these value judgements is clearly spelled out.
- It is difficult to decide on the most appropriate model until the analysis of the supporting data can be performed. Our process delays a commitment to a specific model until the quality of the available data for all of the models is evaluated.

2 DESCRIPTION OF THE PROPOSED APPROACH

2.1 General Approach

The general approach we propose and the various stages involved are outlined below.

Step 1: Several strategies (models) for assessing risk are predefined. Examples of such models are given in section 1.2.

Step 2: The assumptions for each model are defined. For example, one of the assumptions for 1.2.1 described in section 1.2.1 states that the risks from the exposure to the ingredients of the mixture are additive.

The assumptions are often shared by different models, but each model is based by unique set of assumptions.

Step 3: Each assumption is evaluated and the validity of each is determined. The validity of the assumption is expressed as a numerical rank.

It is presumed, that the robustness of a model is a function of the validity of the assumptions upon which the model is based. For many assumptions, the answer is not a simple "valid" or "not valid". The confidence in the available information may vary based on quality, quantity and general agreement within the literature. The assumption may also be valid under some, but not other conditions (limits). The assumption may also not be strictly speaking valid, but it may approximate reality closely enough to be treated as valid.

Step 4: The robustness of each of the models is ranked.

The ranking takes into account the validity of the assumptions on which the models are based and the dependence of the models on individual assumptions.

Step 5: Two or more of the highest ranked models may be used for the evaluation of risk. The results from these risk assessments can be compared and the best estimate of risk can be recommended for the development of a standard or guideline.

2.2 Application of the approach to PAH and PAH-containing mixtures

2.2.1 Description of the models

In section 1.2 we have described the different models used by different regulatory agencies and researchers to estimate the risk to man from exposure to PAH. In this section, We will present the approaches we are considering for the dose-response assessment of PAH.

2.2.1.1 Individual Compound Model

The risk associated with any given mixture is then estimated from the sum of the individual risks contributed by each PAH in the mixture. In order to use the Individual Compound Model, a quantitative chemical analysis is required for each mixture or medium for which a quantitative health risk is to be established.

Since human data is not available for individual PAH, the activity of each PAH is determined from experimental long and short-term data.

2.2.1.2 Grouped Compound Model

Individual PAH are grouped into a few classes, which reflect similarities in structure, relative activity and assumed similarities in the mechanisms in action. Each group is "assigned" a potency. Every compound within a group is assumed to have the same potency. The potencies of the groups, rather than individual PAH are summed in order to estimate the activity of the mixture as a whole.

2.2.1.3 Lumped Compound Model

All PAH-containing mixtures, such as coke oven emissions, are "assigned" the same potency. The risk estimate for a particular mixture is based on some measure of its concentration in a given medium and the assigned potency. In essence, this approach treats all mixtures as "essentially same" and consisting of a "single compound".

2.2.1.4 Models based on equivalents of single compound or mixture

2.2.1.4.1 B[a]P equivalents.

The potency of each individual PAH is expressed in terms of equivalents of a "classical" PAH (such as B[a]P). In order to estimate the activity of a particular mixture, the quantities of each PAH in the mixture is determined, translated into the number of B[a]P equivalents and summed up. This approach is numerically equivalent to the approach outlined in section 2.2.1.1. This approach is advocated by Thorslund and Charnley, 1988 [1].

2.2.1.4.2 Coke oven emission or Cigarette smoke condensate equivalents

The activity of each mixture is converted into the equivalents of a "standard" mixture with a known risk (Coke oven emissions or cigarette smoke condensate could be used as standards)

2.2.1.5 Source Mixtures Model

Unit risk is estimated for each type of source (coke oven emissions, diesel emissions etc.). The relative contribution of risk of each of these Source Mixtures to the overall mixture at the receptor level (an area where target is exposed, called Receptor Mixture) is estimated. The expected carcinogenic risk of the Receptor Mixture is calculated by summing the risks attributable to each of the sources (see equation 3:

Equation 3.

$$TCR = CM1 + CM2 + CM3 + \dots CMn = (CS1 * QS1) + (CS2 * QS2) + \dots$$

where:

- TCR = total carcinogenic risk of all Receptor Mixtures derived from all media (air, ambient water, drinking water...).
- CM1, CM2 = total carcinogenic risk of all Receptor Mixtures derived from a specific medium.
- CMn = total carcinogenic risk of all Receptor Mixtures derived from nth medium.
- CS1, CS2 = carcinogenic activity of a particular type of PAH source per unit quantity
- QS1, QS2 = quantity of a particular type of PAH source

2.2.1.6 The Monitoring Model

The carcinogenicity of an environmental mixture containing PAH is estimated from the genotoxicity of this mixture in short-term assays.

2.2.1.7 The "Blending Model"

The Blending approach simulates blending of the mixtures for which the risk and the composition is known (known mixtures) to reconstitute a mixture with the same composition as the mixture for which the risk is to be estimated (unknown mixture). The risk associated with the unknown mixture is estimated by summing the risk contributed by each of the known mixtures. In some cases, it may become necessary to add some individual PAH to the blend in addition to the known mixtures. It should be stressed however, that it is not assumed that the unknown mixture was actually formed from the known mixtures as calculated in the Blending approach.

2.2.2 Identification of the assumptions

Each of the models described above is based on a series of assumptions. The most robust model is the model based on the most robust set of assumptions. None of the models requires all of the assumptions listed in table 2. Each model is based on a different mix of assumptions from that table.

It is not expected that any of the models will satisfy all of the assumptions completely.

Collectively the assumptions required for any given model indicate how appropriate (see section 1.3) the data set is for that model. In an ideal model, where the risk could be determined directly on the population of interest, few assumptions would be required and the data set would therefore contain the most appropriate data set.

2.2.3 Ranking of the models

A robust model is based on valid assumptions. In the real world, however, it is difficult to categorically state, whether an assumption is valid or not. Often, there is not enough high quality data to evaluate an assumption with confidence. An assumption may also be valid or not valid, depending on the exact circumstances under which the assumption is applied. For example, an assumption (no. 2 in table 2) that there is a quantitative correlation between long-term and short-term animal data may hold for nitrosamines, but not for compounds in general [21]. A given assumption may also only approximate the experimentally derived data. For example, an assumption of a linear dose-effect relationship may not be correct at high doses, but on the whole the linear estimation may still offer a reasonable approximation of experimentally derived data. In practice therefore, the models typically depend on assumptions that can be only partially validated and which may hold only imperfectly and under only some circumstances. The validity of a given assumption is expressed as the **Validity Score (V)**. The process of deriving the score is outlined in section 2.2.3.1.

The robustness of a given model depends, not only on the validity of its assumptions, but also on the degree to which that model is dependent on each of the assumptions. For example, both the Individual Compound Model and the Monitoring Model depend on assumption #2 from table 2 (Strong correlation between long-term and short-term animal data). Since monitoring is practical only with short-term tests, the Monitoring Model is very dependent on the assumption that short-term tests can predict the potency in the long-term tests. In contrast Individual Compound Model can utilize results from both the long-term and short-term tests, and thus, is less dependent on short-term tests. As a result, the Monitoring Model is much more dependent on the assumption #2 than the individual Compound Model. The dependence of a model on a given assumption is expressed as a **Dependence Score (D)**.

The degree of dependence of a model on its assumption is a property of that model and is not affected by the nature of a compound or mixture (data), to which the model is applied. As a result, it is not necessary or appropriate to change a dependence score when new sets of data are applied to the model. In contrast, the validity of a given assumption is based strictly on the supporting data.

2.2.3.1 Ranking Validity of assumptions

The Validity of an assumption ranges between - 10 points and 0 points. In effect, points are subtracted from 0 ("ideal") to account for deficiencies in the validity of the assumptions.

The score is determined as follows:

For each assumption, the points for each of the applicable statements listed below are summed. A non-applicable statement has a score of 0 points. If there

Table 2. List of assumptions for the Data Models

No	Assumption
1	Strong correlation between long-term animal data and human data
2	Strong correlation between long-term and short-term animal data
3	PAH contribute a known proportion of an overall activity of a mixture
4	Carcinogenic activity of all organic substances in a mixture is known
5	Carcinogenic activity of a mixture as a whole is known
6	Carcinogenic activity of individual components of the mixture is additive
7	Carcinogenic activity of individual mixtures is additive
8	Detailed analytical data for all Source Mixtures are available
9	Detailed analytical data for all Receptor Mixtures are available
10	A mixture from a given source always has the same composition
11	Carcinogenic activity of all mixtures is the same
12	Composition of a mixture at the source and receptor level is the same
13	Source markers for all source mixtures are available
14	Source markers form a fixed proportion of Source and Receptor mixtures
15	Transformation markers are a reliable measure of PAH environmental transformation.
16	It is possible to predict the bioavailability of particle-bound mixture components

are no data to evaluate the statement, the average of the range is assumed. The assumptions are evaluated on three validity scales: Confidence, Limits and Accuracy. The scoring using these scales is described in the table 3.

Table 3. Scheme for scoring validity of assumptions

Rule no	Category	Score Range	Score Description
1.a	Confidence (Quality and availability of data.)	0 to - 3	Not enough reports applicable to test an assumption. Key information is not available.
1.b		0 to - 2	Inconsistencies in the results or poor quality of the data.
2.a	Limits (Validity under defined conditions.)	0 to - 1	The assumption may hold only within specified limits.
2.b		0 to - 1	There is evidence, that the assumption does not hold under the "real world" conditions under which the models may be applied.
3.a	Accuracy (assumption "approximates" real world situation)	0 to - 1	The assumption involves a relation between the two or more factors (such as dose and response) and the relationship between these factors is not well known.
3.b		0 to - 2	The assumption involves a relation between the two or more factors (such as dose and response) and the relationship between these factors is known and the assumption poorly approximates the relationship within the relevant range.

2.2.3.2 Dependence of models on assumptions

The scoring scheme for the dependence of a model on its assumptions is described below. A score between 0 and 10 is assigned based on the scoring scheme described in table 4.

Table 4. Scheme for scoring dependence of the models on assumptions

Score Range	Score Description
0	Model is independent of a given assumption
1 to 3	Model is somewhat dependent on the assumption, but it is unlikely, that the validity of the assumption will have decisive effect on the robustness of the model.
4 to 7	Model is clearly dependent on the assumption. When the Validity of the assumption is low, the dependence of this model on this assumption counts against it.
7 to 10	Model is critically dependent on the assumption.

Table 5. lists the proposed Dependence Score for three sample models. The same approach is applicable to the models described in section 1.2 or other models defined by a user of this approach. Note that the dependence score provides a convenient summary of the value judgements made during selection of data sets. The other component of the value judgement summary is the set of validity scores for the individual assumptions.

Table 5. Proposed Dependence of models on their assumptions

Assumption Number	Dependence Score		
	Lumped Compound	Sources Mixture	Monitoring Model
1	0	3	10
2	0	1	10
3	0	0	0
4	0	0	0
5	7	10	0
6	0	1	0
7	0	8	0
8	0	8	0
9	0	8	0
10	0	8	0
11	7	0	0
12	0	4	0
13	0	4	0
14	0	4	0
15	0	2	0
16	1	2	10

2.2.3.3 Ranking of robustness of the models

The Validity Scores and Dependence Scores are used to rank the models. Two ranking schemes have been developed (General Rank and Focused Rank). It is advisable to determine both the General and Focused Rank for each model, in order to identify all models which are not suitable for risk assessment of a given group of chemicals.

2.2.3.3.1 Determination of General Rank

Some models are complicated and depend on a number of assumptions. Even if the assumptions are relatively valid, their number tends to reduce the confidence in a model. Source mixture model is an example of such a model (see table 5). There is only one assumption with a Dependence Score of "10", but few assumptions with a score of "0". The General Rank (Rg) is most applicable to such a model.

The General Rank Rg for a given model is determined by summing up the products of the Validity of a given assumption and the Dependence of the model on that assumption (see the equation 4 below).

Equation 4.

$$Rg = \sum_{a=0}^n V_a * D_a$$

where:

- n is the number of assumptions considered for a given model
 V_a is the Validity of ath assumption
 D_a is the Dependence of the model on ath assumption
R is a non - positive number, since V_a is a non - positive number ranging between 0 and - 10 and D_a is a non - negative number ranging between 0 and 10. The more negative the Rank the less robust is the model. Ideally robust model has a Rank of 0.

2.2.3.3.2 Focused Rank

Some models may be dependent on only a few assumptions, but the dependence on these assumptions may be considerable. If the General Rank was used, the model may not fail, because the failings of a few assumptions is offset by independence from other assumptions. Therefore a Focused Rank

(Rf) is used. An example of such a model is the Monitoring Model. The Dependence Scores are "0" for most assumptions, but for a few, the score is "10", indicating heavy dependence.

Focused Rank is determined by summing the Product Score for each assumption. The Product Score for any given assumption is the score assigned to the product of Validity and Dependence for that assumption (see table 6). For example, let us assume, that the Validity of an assumption is - 8 and the Dependence of the model on that assumption is 4. The product of the two numbers is $(-8) * 4 = -32$. Table 6 below shows that the Product Score for product between - 30 and - 39 is 1. By summing Product Score for all 16 assumptions, the Focused Rank is determined.

Table 6. Derivation of Product Score from the product of Validity and Dependence.

Product	Product Score
0 to - 29	0
- 30 to - 39	1
- 40 to - 49	3
- 50 to - 59	7
- 60 to - 69	12
- 70 to - 79	18
- 80 to - 100	reject model

2.2.3.4 Selection of models

The models with the least negative score in both the General Rank test and the Focused Rank test are the models that are judged to give the most reliable risk assessment. Some tests perform well in one of the tests but poorly in the other. It is the poorer of the two performances which should be given a higher weighting in prioritizing the models. We do not recommend using the results of the two ranking tests in a mechanical manner. The final selection calls for expert judgement and the results from the two ranking tests are a tool to aid in the selection.

3 THE SIGNIFICANCE AND APPLICATION OF OUR APPROACH

Our approach has several useful features. That include:

3.1 Facilitation of literature gathering and analysis

One of the most important features of the described approach is its usefulness in the gathering and analysis of the literature. The US EPA guidelines provide general

direction for the types of material that ought to be considered during risk assessment on mixtures. The guidelines do not, however, provide sufficient detail or specificity to direct the literature analysis. In contrast, the described approach defines the specific tasks required to perform the literature analysis in some detail by defining assumptions which need to be tested. Our process thus directs the literature analysis and provides a checklist which assures that the coverage was adequate.

3.2 Formal and simple process for selection of model (s) for subsequent risk assessment

Our approach permits direct comparison of robustness of the approaches to risk assessment. Many different factors are taken into consideration at the same time. In contrast, the US EPA recommends the risk assessment based on mixtures, whenever possible. The US EPA concedes, that under some circumstances, risk assessment based on the individual components should also be considered, but the principles on which to base these value judgements are not well defined.

3.3 Self - documentation of the model selection process.

Performance of the risk assessment calls for frequent value judgements. The results of the assessment are often significantly dependent on the nature of the judgements made. Our approach documents the key value judgements as the Dependency Score and the Validity Score. The reviewer of a document which uses our strategy is thus in a much better position to find what value judgements were made. The US EPA guidelines provide no such feature.

3.4 Facilitated risk assessment process.

Many components of the model selection process are directly reusable and applicable in the risk assessment process. This significantly reduces the "overhead" for the risk selection process.

3.5 Objective evaluation of available options

The structure of the proposed process helps reduce the impact of the personal biases of the investigator (s).

3.6 Versatility - applicable to single compounds, mixtures, exposure assessment

The key activities of the described approach are:

- o Define Models
- o Compile the list of assumptions for these models
- o Generate Dependence Scores for all models and all assumptions
- o Generate Validity Scores for all assumptions
- o Rank the models based on the Dependence and Validity Scores.

This approach can be applied to single compounds as well as complex mixtures and it can be applied to the exposure assessment modeling as well.

3.7 Modular - facilitates team effort

The bulk of the work in this process involves evaluating the validity of assumptions. Evaluation of each assumption represents a discreet task which can be assign to different members of the team. This approach lends itself to teamwork well

4 APPLICATION OF THE APPROACH TO A SET OF DATA - AN EXAMPLE.

In order to apply the approach to any specific set of data, the scoring of the assumptions must be completed. Our group is not yet ready to do this scoring. For demonstration purposes only, a set of scores has been created and presented in table 6. THESE SCORES DO NOT REPRESENT THE MINISTRY'S SCORES FOR THE PAH AND THE PAH-CONTAINING MIXTURES. The purpose of this exercise is to illustrate the process and demonstrate the interpretations possible with this approach.

The steps involved in ranking the models are described below:

4.1 Define models and their dependence on assumptions

The models are described in section 1.2 and the Dependence Scores are summarized in table 5.

4.2 Determine the Validity Score for the assumptions

A simulated Validity Scores were derived for each assumption (see table 2) using the scoring scheme defined in table 2. An example of the derivation of the Validity Score is presented in table 7. The individual scores under each scoring category (table 7) are summed up to give the Validity Score. The validity score was simulated for the other assumptions in a similar way as the Validity Scores are presented in table 7.

Table 7. Derivation of simulated Validity Scores for assumption #2.

Category	1.a	1.b	2.a	2.b	3.a	3.b	Validity Score
Score	-2	-1	-1	-1	0	-2	-7

Table 8. Simulated Validity Scores for assumptions.

Assumption number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Validity Score	-3	-7	-4	-5	-3	-7	-3	-3	-3	-8	-4	-3	-7	-5	-5	-4

4.3 Determine the products of Validity Scores And Dependence Scores for each model.

The products (Pg) of Validity Scores (Va) and Dependence Scores (Da) are required to calculate both the General Rank and the Focused Rank. The products were obtained as described in section 2.2.3.3.1. and summarized in table 9.

4.4 Determine the General Rank

In order to calculate the General Rank, the products for each assumption in the model were summed up. General Rank score for three sample models is presented in the "Rg" row of table 9.

4.5 Determine the Focused Rank

In order to calculate the Focused Rank, the products (Pg) were first converted to Product Scores using table 6. The Product Scores for each assumption in the model were then summed up, in order to obtain Focused Rank. Focused Rank score for three sample models is presented in the "Rf" row of table 10.

4.6 Evaluation of Results

In the specific (simulated) example, the best model is clearly the Lumped Compound model, which performed best of the three models considered in both the General Rank and the Focused Rank Test. It is instructive to look at the differences between the performance of the models in the two ranking schemes. The Source Mixture model is dependent on a large number of assumptions, but the assumptions are either relatively valid, or the model has low dependence on them. This model received the worst score in the General Rank test, because of its dependence on large number of assumptions. In contrast, the model did well in the Focused Rank test, because of its low dependence on any of the assumptions. In contrast, Monitoring model is dependent critically on a

few assumptions, which, in our simulation were not very valid. As a result, it performs better in the General Rank test than in the Focused Rank test, even though it does poorly in the Focused Rank test.

In general, both the Focused Rank and the General Rank answer different questions and should be used in parallel.

Table 9. Dependence Scores (Da) and Simulated Validity Scores (Va) and their products (Pg) for three sample models. The table also provides the General Rank (Rg) and the Focused Rank (Rf) for the three models

Ass.	Va	Lumped Mixtures		Source Mixture		Monitor.	
		Da	Pg	Da	Pg	Da	Pg
1	-3	0	0	3	-9	10	-30
2	-7	0	0	1	-7	10	-70
3	-4	0	0	0	0	0	0
4	-5	0	0	0	0	0	0
5	-3	7	-21	10	-30	0	0
6	-7	0	0	1	-7	0	0
7	-3	0	0	10	-30	0	0
8	-3	0	0	8	-24	0	0
9	-3	0	0	8	-24	0	0
10	-8	0	0	7	-56	0	0
11	-4	7	-28	0	0	0	0
12	-3	0	0	4	-12	0	0
13	-7	0	0	4	-28	0	0
14	-5	0	0	4	-20	0	0
15	-5	0	0	2	-10	0	0
16	-4	1	-4	2	-8	10	-40
Rg			-53		-265		-140
Rf			0		9		22

Table 10. Simulated General and Focused Ranks (Rg and Rf respectively) for different models. Order of preference based on each rank is given in the columns labeled "Order".

Model	General Rank		Focused Rank	
	Rg	Order	Rf	Order
Lumped mixtures	-53	1	0	1
Source mixtures	-265	3	9	2
Monitoring	-140	2	22	3

References

1. Thorslund, T. W.; Charnley, G. 1988
Comparative potency approach for estimating the cancer risk associated with exposure to mixtures of polycyclic aromatic hydrocarbons.
ICF-Clement Associates, Fairfax, Virginia.
2. Albert, R. E., Lewtas, J., Nesnow, S., Thorslund, T.W., Anderson, E., 1983
Comparative potency Method for cancer risk assessment: Application to Diesel Particulate Emissions.
Risk Analysis, 3 ; 101-117
3. Ontario Ministry of the Environment, 1985
Polychlorinated dibenzo-p-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs)
Scientific Criteria Document For Standard Development No. 4-84
4. US EPA (1984)
Health effects assessment for polycyclic aromatic hydrocarbons (PAHs)
EPA 540/1-86-013. Environmental Criteria and Assessment Office, Cincinnati.
5. John Summerhays, 1989
Estimation and evaluation of cancer risks attributable to air pollution in south-east Chicago.
US EPA, Air and Radiation Division, Chicago, Ill
6. Pike, M.C., Henderson, B.E., 1981
Epidemiology of polycyclic hydrocarbons: Quantifying the cancer risk from cigarette smoking and air pollution
Polycyclic hydrocarbons and cancer, vol. 3, Gelboin, H.E. And Ts'o, P.O.P. Editors. Academic Press, pp 317 - 333
7. WHO, 1984
Guidelines for drinking water quality. Vol.2
pp 182-189
8. WHO 1987
Air quality guidelines for Europe
WHO regional publications, European series no. 23
pp 105-117

9. Anderson, E.L.; Carcinogen Assessment Group of the U.S. Environmental Protection Agency, 1983
Quantitative approaches in use to assess cancer risk.
Risk Analysis **3**, 277 - 295
10. Krewski, D.; Birkwood, P.L., 1987
Risk Assessment and Risk Management
Risk Abstracts, **4**, pp 53 - 61
11. Wilkinson, C.F.; 1987
Being more realistic about chemical carcinogenesis.
Environmental Science and Technology **21**, pp 843 - 847
12. Karstadt, M.; 1988
Quantitative Risk Assessment: Qualms and Questions.
Teratogenesis, Carcinogenesis and Mutagenesis **8**, pp 137 - 152
13. Rodricks, J.V.; Tardiff, R.G.; 1984
Conceptual Basis For Risk Assessment.
Assessment and Management of Chemical Risks. ACS Symposium Series 239;
Rodricks, J.V. And Tardiff, R. G. Editors
American Chemical Society, pp 5 - 12
14. Brown, C. C.; 1984
High- to Low-dose Extrapolation in Animals.
Assessment and Management of Chemical Risks. ACS Symposium Series 239;
Rodricks, J.V. And Tardiff, R. G. Editors
American Chemical Society, pp 57 - 79
15. Wilson R.; Crouch, E.A.C.; 1987
Risk Assessment and Comparisons: An Introduction.
Science **236**, pp 267 - 270
16. Nelson, N.; 1984
Use of Toxicity Test Data in the Estimation of Risks to Human Health.
Assessment and Management of Chemical Risks. ACS Symposium Series 239;
Rodricks, J.V. And Tardiff, R. G. Editors
American Chemical Society, pp 13 - 22
17. Menzel, D. B.; Smolko, E.D.; 1984
Interspecies Extrapolation.
Assessment and Management of Chemical Risks. ACS Symposium Series 239;
Rodricks, J.V. And Tardiff, R. G. Editors
American Chemical Society, pp 23 - 35

18. Visek, W.J.; 1988
Issues and Current Applications of Interspecies Extrapolation of Carcinogenic Potency as a Component of Risk Assessment.
Environmental Health Perspectives **77**, pp 49 - 54
19. US EPA, 1986
Guidelines for carcinogen risk assessment
Federal Register **51**, pp 33992-34003
20. US EPA; 1986
Guidelines for the Health Risk Assessment of Chemical Mixtures.
Federal Register **51**, pp 34014-34025
21. Piegorsch, W.W.; Hoel, D.G.; 1988
Exploring relationships between mutagenic and carcinogenic potencies.
Mutation Research **196**; pp 166 - 175



(7801)

TD/5/T43